# Managing Stress: The Needs of Autistic Adults in Video Calling

ANNUSKA ZOLYOMI, Microsoft Research, University of Washington
ANDREW BEGEL, Microsoft Research
JENNIFER FRANCES WALDERN, Microsoft
JOHN TANG, MIKE BARNETT, ED CUTRELL, DANIEL MCDUFF, SEAN ANDRIST, and MEREDITH RINGEL MORRIS, Microsoft Research

Video calling (VC) aims to create multi-modal, collaborative environments that are "just like being there." However, we found that autistic individuals, who exhibit atypical social and cognitive processing, may not share this goal. We interviewed autistic adults about their perceptions of VC compared to other computer-mediated communications (CMC) and face-to-face interactions. We developed a neurodiversity-sensitive model of CMC that describes how stressors such as sensory sensitivities, cognitive load, and anxiety, contribute to their preferences for CMC channels. We learned that they apply significant effort to construct coping strategies to support their sensory, cognitive, and social needs. These strategies include moderating their sensory inputs, creating mental models of conversation partners, and attempting to mask their autism by adopting neurotypical behaviors. Without effective strategies, interviewees experience more stress, have less capacity to interpret verbal and non-verbal cues, and feel less empowered to participate. Our findings reveal critical needs for autistic users. We suggest design opportunities to support their ability to comfortably use VC, and in doing so, point the way towards making VC more comfortable for all.

CCS Concepts: • **Human-centered computing → Computer supported cooperative work**; **Empirical studies in accessibility**.

Additional Key Words and Phrases: video calling, computer-mediated communication, accessibility, autism, neurodiversity

**ACM Reference Format:**
Annuska Zolyomi, Andrew Begel, Jennifer Frances Waldern, John Tang, Mike Barnett, Ed Cutrell, Daniel McDuff, Sean Andrist, and Meredith Ringel Morris. 2019. Managing Stress: The Needs of Autistic Adults in Video Calling. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article X (November 2019), 29 pages. https://doi.org/10.1145/TBD

## 1 INTRODUCTION

Video calling (VC) applications provide multi-modal environments that offer real-time video, audio, chat, and desktop sharing channels to create common ground for collaboration in many domains, including work, school, and personal scenarios [56]. Video calls and other computer-mediated

Authors' addresses: Annuska Zolyomi, annuska@uw.edu, Microsoft Research, University of Washington, Seattle, WA; Andrew Begel, abegel@microsoft.com, Microsoft Research, Redmond, WA; Jennifer Frances Waldern, Jennifer.Waldern@microsoft.com, Microsoft; John Tang, johntang@microsoft.com; Mike Barnett, mbarnett@microsoft.com; Ed Cutrell, cutrell@microsoft.com; Daniel McDuff, damcduff@microsoft.com; Sean Andrist, sandrist@microsoft.com; Meredith Ringel Morris, merrie@microsoft.com, Microsoft Research, Redmond, WA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2019 Association for Computing Machinery.
2573-0142/2019/11-ARTX $15.00
https://doi.org/10.1145/TBD

communication (CMC) tools connect people across distance, often when they would otherwise be unable to meet face-to-face (FtF). These tools can be hyper-personal—enabling the senders, receivers, communication channels, and messages to work together to bolster interpersonal relationships [53].

Some people who want to use VC are on the autism spectrum. Autism spectrum disorder (ASD) is a lifelong neurodevelopmental condition characterized by particular cognitive styles, communication behaviors, social interactions, and repetitive behaviors [3]. Recent estimates put the prevalence of autism in the USA at 1 in 42 in boys and 1 in 189 in girls [4, 5], a rate which has risen drastically over the past 30 years as awareness has increased and diagnosis processes evolve. As autistic individuals[1] reach adulthood and enter the workforce, new VC scenarios and user requirements are likely to arise to support collaboration between autistic and non-autistic (i.e., neurotypical[2]) collaborators.

During VC interactions, people's goals range from building social relationships to collaborating towards completing a group task. While attending to these meta-cognitive goals, each person has to draw upon their low-level cognitive processing capabilities to manage the multi-modal channels of audio, video, text, and images. These processing demands can be challenging for neurotypical users because of natural limitations on human processing abilities [15, 37]. As a result, researchers have examined ways to help VC users more efficiently manage meta-cognitive and low-level cognitive processing (e.g., [16, 44, 56]).

Socio-technical challenges are more intense for many autistic VC users. Due to current social norms and the design of VC systems, autistic individuals have to adapt their particular style of cognitive processing and hyper- or hypo- sensitivity to sensory inputs during a VC. Their impressions of using VC are likely impacted by how they cognitively process VC interactions. Autistic individuals report having difficulty expressing their emotions, adapting to new situations, and working through ambiguity [19, 22]; they tend to focus on details and excel at conceptualizing phenomena as systems [6]. During VC situations, their sensory sensitivities may be irritated, especially when lights prove to be too bright and microphone noises sound too loud and are distracting [19, 22]. The high-bandwidth communication channels offered by video calls may easily trigger an autistic individual's detail-focused cognitive style to devote too much time, attention, and effort trying to read others' emotions and body language, causing them to fall behind in conversational flow [22, 42]. Though VC affords "some of the intimacies of co-presence," [23, p. 68], the desire for that affordance may not be shared by autistic users.

Unfortunately, little is known about the strategies and technology decisions of autistic adults as they engage (or disengage) with VC. In this paper, we aim to fill this gap by asking

RQ1. How and why do autistic adults make use of video calling?
RQ2. What factors increase or reduce the comfort of the video calling experiences of autistic adults?
RQ3. How do these comfort-influencing factors impact the CMC channel preferences of autistic adults?
RQ4. What coping strategies do autistic adults make use of to reduce discomfort and relieve stress during video calls?
RQ5. How could video calling tools be changed to better accommodate autistic users?

To answer these questions, we conducted 22 semi-structured interviews with autistic adults to learn about their perceptions of the benefits and drawbacks of VC. We found that they experienced difficulties with technology-mediated social norms at every stage of VC, from preparing for calls,

[1] While some in the autism community prefer people-first language, others have embraced the term "autistic" as their chosen identifying label, so we use that terminology as well [29].
[2] According to the neurodiversity framing of autism, there is a natural diversity of human neurological functioning, including autistic (neurodiverse) and non-autistic (neurotypical) cognitive styles [47].

initiating calls, conversing with others, and ending calls. These stressors drove them to use VC technologies in unexpected ways to improve their comfort levels. We found that they employed coping strategies to adapt their VC environment and their behaviors to support their sensory, cognitive, and social needs during a call. These strategies included (1) controlling their sensory experiences to improve their focus, (2) using strategies, such as writing notes, to retain relevant information, and (3) developing a clear mental model of their conversation partner's affect and cognitive style. When lacking appropriate strategies, interviewees reported becoming more stressed, less able to interpret social-emotional cues, and less effective in their role for the meeting. To describe the relationship between stressors, coping strategies, and CMC preferences, we developed a neurodiversity-sensitive model of computer-mediated communication (NDS-CMC) (see Section 5.1).

Studying VC through the lens of autistic users enabled us to discover rich user needs that were immediately evident, and potentially salient for neurotypical users, as well. For example, conversational dynamics, such as turn-taking and knowing when to end a meeting, can often feel ambiguous. Social-emotional cues are easily misinterpreted, especially when a person's words are incongruent with their voice tone and body language. In Section 6, we reflect on how we can apply the lessons we learned about the VC experiences of autistic adults to suggest ways to design VC and CMC tools to better support the sensory, cognitive, emotional, and social needs of all of its users, autistic and neurotypical.

Many researchers have discovered that interventions for people with disabilities can help the greater population and support this notion in their work. Picard has learned and applied this regarding the benefits of software that supports affect awareness [40]. Burke et al. has similarly extended this lesson as applied to CMC-specific social skills training for autistic adults [12]. We believe the lessons we learned in this study can help point the way towards features that could make VC more comfortable for everyone.

To summarize, in this paper, we make three contributions:

- We identify how and why autistic users use VC and other CMC channels (Section 4.1).
- We introduce a neurodiversity-sensitive model of CMC that explains the stressors, CMC preferences, and coping strategies of autistic individuals making video calls (Sections 5.1, 5.2, and 5.3).
- We suggest ways to make VC more comfortable for autistic users and their conversation partners (Section 6).

## 2 BACKGROUND

Our research builds on trends in VC research for mainstream audiences. We provide a background to autism spectrum disorder and introduce several theories that may explain a subset of the difficulties our study interviewees revealed as they made use of VC. From there, we consider current knowledge of the perceptions of autistic adults using the Internet and some forms of CMC. Finally, we summarize the gap between prior literature and what we investigated in our study.

### 2.1 Video Calling

Video calling comes out of a long history of computer-supported cooperative work (CSCW) systems that enable people to connect across distances [21]. VC platforms bring participants together into a common virtual space using the modalities of live video, visual sharing of the desktop screen, an auditory channel, and text chat. When choosing a communication channel for collaborating and communicating with others, CMC theories and models highlight the channel affordances, the interaction partner, and the interaction topic [17, 53].

VC can provide a rich experience because it offers a range of channels within one system, thus enabling users to leverage the affordances of multiple channels, e.g. video, audio, and/or text, among others. Media Richness theory helps explain how people choose CMC channels based on each channel's set of objective characteristics that determines its capacity to carry rich information [18]. In this context, information is rich if it helps the sender and receiver to communicate clearly and adjust their understanding as necessary. According to the theory, VC collaborators have a greater likelihood of conducting clear, unambiguous communications using audio and video than if they were communicating over a leaner channel (e.g. text) which would afford fewer sensory modalities. In addition to a channel's ability to carry rich information, users also chose channels based on how much control they have over them, their ability to remain anonymous if desired, and the ability to be co-present [17]. The notion of co-presence and distance is complicated by the collaborators' sense of distance; it is not as straightforward as whether or not the collaborators are literally in the same room. Research has looked at the effects of perception of distance and social context during VC. For example, Bradner and Mark found that VC collaborators exhibited more positive collaborative behaviors (e.g., cooperation) when they believed their collaborator was nearby, rather than far away [11].

Along with distance, other important attributes of the interaction partner that impact technology-mediated collaboration are social ties and the social factors relating to the context of the topic of conversation [17]. When text-based CMC became prominent, the social information processing (SIP) model [52] and hyper-personal model of CMC emphasized that users' experiences with CSCW were influenced by the social relationships among the participants. According to the hyper-personal model, users' social ties can become stronger through the coordination of four concurrent CMC *routines:* actions of the receivers, senders, channel, and continual feedback among those three components [53]. As these routines reinforce each other they can facilitate greater social desirability and intimacy that are of a different nature (and sometimes better) than developed in FtF interactions. This theory argues that attributes of CMC, such as temporal features, influence its capacity to build social relationships. For example, media preferences for populations that tend to face challenges of real-time FtF communication, such as people who are non-native speakers, often prefer text channels so that they can have more time to process messages and craft responses [46]. Although the SIP model and hyper-personal model focused on text channels, researchers have continued exploring social factors in CMC by extending the models to other modalities (e.g., online gaming [54]) and by conducting empirical research based on other CMC theories. The Embodied Social Proxy research explored a mobile VC terminal to enable interacting with a team in the context of social meeting spaces [51]. This approach evoked more social interactions that improved the remote collaborator's social integration with the team. How video technology is appropriated to reflect the context of social relationships can become more salient in a neurodiverse population that is sensitive to managing or controlling their sense of distance with their conversation partners. Our work is informed by this line of research, as we consider how autistic adults may experience strengthened social relationships using CMC channels that may be preferred over FtF interactions.

While VC naturally affords transmitting visual and nonverbal cues exchanged in an interaction, research has also begun to explore computationally detecting and representing those cues. Byun et al. [13] explored analyzing audio and video streams in a video call to detect visual and non-verbal cues and display them in real time to the interviewees to make their calls more successful. They found that users appreciated the feedback to help them manage their VC behavior. Grayson and Monk examined the establishment of mutual eye contact through VC, and found that users are able to learn to interpret eye gaze direction in VC, even without perfect mutual eye contact [25]. Beyond eye contact, facial expressions convey affective information and provide socio-emotional feedback. Facial expressions are particularly helpful in interpreting the affective state of users, illustrated

by Ekman's work showing that people across a number of different cultures can recognize seven distinct emotions from photos of facial expressions [20]. These explorations into real-time feedback and adaptations suggest interesting affordances of VC for a neurodiverse population who experience difficulties with non-verbal cues and with expressing and interpreting emotions.

Our study examines the social dynamics involved in using CMC, plus the perceptions of VC compared to other CMC. Our ability to take another person's perspective increases when we can read one another's body language and engage in a micro-level exchange of non-verbal interactions that help one form an opinion about another person and what they are communicating. In our research, media richness theory suggests that autistic adults will prefer VC over other forms of CMC because the richness of the VC channel will provide them with the multi-modal experience to see facial expressions and body language. However, socially-focused CMC theories suggest that autistic users may prefer CMC over FtF, specifically, for building and maintaining relationships because the CMC attributes lend themselves to more flexibility in how conversations are conducted in terms of timing, brevity, etc.

### 2.2 Autism

Scholars in disability studies, along with autistic self-advocates, have articulated three key dimensions of autism: (1) the natural diversity of human neurological functioning, (2) the reality of physical and mental strain people face due to autism and common co-morbid conditions, such as gross motor impairments and depression, and (3) the difficulties they face because of the social and physical barriers in our neurotypical world [47]. Common characteristics of autism can be broadly categorized as differences in cognitive styles, having a different sense of social interactions and communication, and exhibiting restricted physical behaviors and interests [3]. Generally, autistic individuals exhibit average to superior abilities to conceptualize abstract and concrete phenomena as systems, including numeric, natural, mechanical, and social systems [2]. Baron-Cohen calls this a hyper-systemizing cognitive style [6], which benefits the individual because it makes phenomena more logical and predictable. An individual with a hyper-systemizing cognitive style has a strong drive to construct and analyze systems, focus on details, and follow rules.

However, this desire to understand the rules, coupled with reported difficulties in taking other people's perspectives, leads to misunderstanding neurotypical social norms that are complex and nuanced. Internal mental states—such as emotions, beliefs, intentions, and attention—are often not readily apparent by one's external behaviors. The ability to attribute mental states to other people is known as a theory of mind, which helps people empathize with others and predict their actions [57]. Some research has demonstrated that, in general, autistic children have difficulty with tasks that rely on theory of mind, such as engaging in pretend play and exhibiting joint attention [7].

Some autistic individuals exhibit repetitive behaviors, including flapping hands, fidgeting with an object, stroking hair, and spinning around. These behaviors can be self-stimulating, in which case these behaviors are referred to as "stimming." One theory about why repetitive motions are important to autistic individuals is the concept of weak central coherence [22, 42]. According to this psychological theory, people exhibit either a strong or weak central coherence style depending on their focus on global or local processing, respectively. The cognitive systems involved in this processing include perception, attention, linguistic, and semantic functions. One theory is that autistic individuals tend to have a weak central coherence style, which implies that they attend to one system at a time. "Mono-channel" attention can make it appear as if they are hyper-focused on a particular topic, sensory input, repetitive noises, etc. Weak central coherence may also impact one's capacity for sharing attention with others and jointly focusing on an object [48].

In the health community, "the spoons theory" provides insight into the socio-emotional work that autistic individuals do to adapt their natural cognitive styles to a neurotypical world [35].

Table 1. Prominent Cognitive Theories of Autism Spectrum Disorder

| Theory | Description |
|---|---|
| Theory of Mind [7, 57] | Ability to attribute mental states to someone else, understand their perspective, and predict their actions |
| Weak Central Coherence [22, 42] | Cognitive style that focuses on the details of situation at the expense of understanding the global context |
| Hyper-Systemizing [6] | Individuals with autism find it useful to view phenomena as a system with inputs, outputs, and rules |
| "The Spoon Theory" [35] | People with chronic health conditions and disabilities spend energy (i.e. metaphorical spoons) throughout the day. Activities can use up or renew spoons. |

"The spoons theory" is the idea that every day, people with chronic health issues or disabilities start with a limited amount of energy (i.e., spoons) and expend it during the day. This metaphor, created by Miserandino, a blogger who lives with lupus, has since propagated into other health and disability communities, including autism. In psychology, this concept aligns with the theories about ego depletion and socio-emotional cognitive resource capacity [8].

In summary, there is no comprehensive theory of autism, and the majority of prominent theories have been critiqued for not evolving from autistic voices. In this paper, we attempt to approach scientific inquiry from the social model of disability [38], which emphasizes that societal barriers, not inherent physical impairments, are the persistent source of obstacles for people with disabilities. Therefore, we critically draw from prominent theories (see Table 1) that have been explored with empirical research to deepen our understanding of common behaviors of autistic people [22]. We also draw from the community-driven "spoon theory" about ego depletion to account for the lived experiences of autistic individuals. Our research is aligned with the disability studies perspective of centering the lived experiences of autistic individuals and the neurodiversity movement [47]. As we consider the intersection of autism and VC, we are sensitive to our interviewees' lived experiences relating to showing emotion, perceiving emotion, and taking the perspective of others. We consider ways that our interviewees' experiences may align with, or contradict, autism theories. For instance, the weak central coherence theory surfaces potential issues with being distracted during VC due to one's physical and sensory environment, conversational partner, and aspects of the VC user interface. The social CMC theories (e.g., the hyper-systemizing model) raise the possibility of autistic adults considering the VC experience as a system, perhaps providing insights into novel ways to improve the experience. Finally, the "the spoon theory" sensitizes us to the socio-emotional work required by VC, leading to our investigation of coping strategies. Although the theories are useful for surfacing cognitive styles and possible explanations for behaviors and perspectives, we remain open to contradictions and new phenomena that our interviews may reveal. We employ a method with flexibility—semi-structured interviews—so we can probe emerging topics.

### 2.3 Computer-Mediated Communication Use by Autistic Individuals

CMC supports communication along many different channels of varied richness. In our research, we consider mainly three channels provided by modern VC tools such as Skype or Google Hangouts: video, audio, and text, as these were the tools and channels most often reported to be used by our study participants. However, as we investigated the context of their choice of CMC channels and their use, we noticed distinct preferences by autistic users for leaner channels than the literature

about CMC use by neurotypical users would suggest. According to Davidson and Henderson's summary of autistic autobiographical authors, *"individuals with ASD generally find face-to-face communication challenging in the extreme"* [19, p. 463]. Burke et al. found that autistic people can feel more at-home using text, email, and social networking CMC channels than interacting FtF [12]. Morris et al. found that among a sample of technology workers, autistic people had higher self-reported levels of comfort with text messaging relative to neurotypical employees, but lower self-reported comfort levels with phone calls, video calls, and FtF conversation in the workplace [36]. Gillespie-Lynch et al. classified the social benefits of CMC (primarily text-oriented channels such as discussion boards, social networking sites, or blogs) for autistic people along two dimensions: (1) "increased comprehension of and control over communication" and (2) "contact with and social support from similar others who may be geographically distant" [24, p. 457]. Note that this classification is aligned with mainstream CMC theorizing that CMC choice is heavily influenced by channel affordances, interaction partners, and interaction topics. Along Gillespie-Lynch et al.'s first dimension of comprehension and control, CMC enables autistic people to engage with other people without the reported strain from experiencing sensory overload and having to deal with the ambiguity of communication [41]. Due to the asynchronous nature of some CMC (e.g., email), autistic users can take time to assess communications that may be ambiguous or contentious. They can formulate their response and follow up with communication partners to clear up miscommunications. Along the second dimension of social contact and support, autistic individuals report valuing using CMC to meet people with similar interests. Platforms dedicated to autistic users, such as the Autcraft Minecraft gaming server, are especially effective socio-technical environments for developing meaningful friendships and sharing experiences with like-minded people [43].

However, as reported by Burke et al. autistic adults reported difficulty in maintaining online relationships due to issues regarding "knowing whom to trust, knowing how much to disclose, and understanding CMC-specific social norms" [12, p. 428]. Compounding these issues is the research insight that people do not always present their full, authentic selves online [27]. Interestingly, autistic adults, more so than neurotypical users, perceive that a benefit of CMC is the opportunity to express their true selves. Alper provides a nuanced and insightful analysis of the implication of technology use in the sensory experiences of autistic children [1]. Audiovisual media content can serve as emulating pro-social sensory seeking activities (e.g., watching animated martial arts movements), yet it can also trigger hyper-sensitivity responses (e.g., repetitive or unexpected loud noises). Alper concludes that "social and technical possibilities converge and diverge around the senses" [1, p. 3573], thus highlighting the need for research in this area. Very little of the ASD literature has looked at experiences or implications of video-based communication between autistic users and their communication partners. Our research fills in these gaps by investigating how autistic adults manage the multi-modal channels of VC and technology-mediated social norms across multiple settings of their work, education, and personal lives.

### 2.4 Literature Gap

In summary, this body of related work raises important areas of inquiry regarding the use of CMC by autistic adults. First, our research can reveal scenarios in which our interviewees' needs (e.g., averting eye contact) do not align with those of neurotypical users. This points to where VCs can build more personalized and inclusive experiences. Second, VCs directly transmit verbal and nonverbal cues and implicitly expect users to interpret them. These skills are far more intuitive for neurotypical individuals than autistic ones. Third, due to the ambiguity and complexity of social norms, an area to explore is the autistic adult's perception of social norms in a VC and their strategies for adapting to the predominant neurotypical norms. Finally, although CMC theories

Table 2. Interview participant demographic information for the 22 interviewees in the study.

| ID | Gender | Occupation |
|---|---|---|
| P01 | M | Software engineer |
| P02 | M | Software engineer |
| P03 | M | Telecommunications engineer |
| P04 | M | Software engineer |
| P05 | M | College student (Computer Repair) |
| P06 | F | ESL teacher |
| P07 | F | High school senior |
| P08 | M | Software engineer |
| P09 | F | Technology Partner Manager |
| P10 | M | Data scientist |
| P11 | M | College graduate (English) |
| P12 | M | College student (Game Design) |
| P13 | F | Healthcare consultant |
| P14 | M | Software engineer |
| P15 | F | College student (Disability Studies) |
| P16 | M | Technology consultant |
| P17 | F | College student (Sociology) |
| P18 | F | College student |
| P19 | M | Analyst |
| P20 | M | College graduate (Information Technology) |
| P21 | F | High school graduate |
| P22 | F | College student (Education) |

present explanations of how people choose CMC channels, an open issue is whether the same conclusions about media richness and the routines of hyper-personal relationships via CMC apply equally to autistic adults.

## 3 METHOD

To investigate our research questions, we conducted a qualitative study by interviewing 22 autistic adults. Our study design was reviewed and approved by our institution's Internal Review Board.

### 3.1 Interview Participants

Interviewees were 18 or more years old and had more than one experience using VC software (e.g. Skype, Google Hangouts, Apple FaceTime, Discord). They were offered a $75 Amazon gift card for participating, whether or not they answered all of our questions. No interviewee stopped early.

Interviewees were recruited through a variety of means. We sent a recruitment email to the authors' own institution's autism mailing lists, one for autistic people and the other for those with autistic relatives. Emails were sent to the disability services offices of all 2-year and 4-year colleges in the authors' USA-based geographical region, asking them to forward the study announcement to their autistic student mailing list. We also advertised with several local area autism service providers and autism support groups. We recruited autistic adults across the USA through a partnership with a private, not-for-profit, rehabilitation hospital on the East Coast. Finally, we advertised online with Asperger Experts, an autism self-advocacy Facebook group.

We interviewed a total of 24 people for the study, but we subsequently analyzed only 22 interviews. We excluded two interviews because the participants had very limited experience with VC and exhibited limited conversational skills, which presented a difficulty in answering our opinion-focused questions about their VC use.

Six interviews were conducted in person, and the remaining were conducted using a variety of CMC tools including phone, Skype, and Google Hangouts. All interviewees live in the USA and range in age from 18 to 49 years old (average 31.5). 13 interviewees identify as male and 9 as female. 12 were employed professionals, six were college students, one was a student in secondary school, and three were unemployed. For more detailed information about the interviewees, see Table 2.

### 3.2 Interview Questions

Interviews provide rich data about a person's opinions and experiences; they also provide opportunities for interviewees to reflect and examine their own behavior. Our interview process was guided by a variety of autism-specific interview strategies described by O'Reilly et al. [39]. Each semi-structured interview was conducted in several parts over the period of one hour and was conducted by two investigators (authors). Both investigators have had extensive prior experience interviewing autistic adults. To minimize distractions, only one investigator interacted with the interviewee; the other mainly took notes, but sporadically asked questions. Interviews were audio-recorded and transcribed professionally for further analysis.

In the interviews, we first requested demographic data, e.g. age, occupation, job role or major. We then asked about interviewees' VC experiences: applications, subject matter, how they prepared for calls, and what they liked and disliked about it. Next, we asked how their autistic characteristics affected their VC experiences, for example asking about subject matter, conversational flow, turn-taking, body language, emotions, sensory sensitivities, and attention. We also asked the interviewees to compare various CMC channels with FtF conversations with neurotypical partners and with others on the autism spectrum. To conclude our interviews, we brainstormed with the interviewee about potential design directions and features to elicit their ideas, reactions, and concerns about ways in which VC could evolve to enhance their VC experiences. [3]

### 3.3 Analysis

We employed an iterative, qualitative style of analysis informed by qualitative research guidelines in Miles and Huberman [34]. Immediately after each interview, the interviewers memoed notes about the compelling experiences, strategies, and emerging concepts from the feature feedback questions. In addition, once a week, the interviewers reported back to the larger research team (i.e., all of the authors) about the highlights from the interviews. The group then engaged in a preliminary level of analysis and brainstorming on design directions. Four researchers conducted more formal analysis by coding the interviews into categories based on the interview topics: VC usage, perceptions of benefits and drawbacks of VC, VC preparation, VC distractions, eye contact and body language, VC compared to other CMC channels, interactions among autistic individuals, and feedback on design concepts. Based on our interview coding, we identified conceptual connections about (1) conversations and emotions, (2) decision factors in choosing VC, and (3) benefits of the various VC channels.

[3]To support replication of our study, the interview script is available at https://github.com/PublicPaperRepository/VideoCallingChallenges/blob/master/Video_Calling_for_Autism_Interview_Questionnaire_Public.pdf.

### 3.4 Limitations of our Research

For this study, we interviewed 22 autistic adults and asked them about their perceptions and personal feelings towards video calls. Some of them lacked experience with various aspects of video calls. In these cases, we asked them about their experiences with phone calls, email, texts, and in-person conversations. These discussions often revealed information about how they perceived their autistic traits to affect their conversations with others, and often affected them independently of any particular CMC channel. Whenever possible, we confirmed with the interviewees whether they were impacted differently between in-person and various CMC channels.

We originally intended to conduct in-person interviews with autistic adults from our local geographical area. However, our efforts did not yield enough recruits, so we expanded our scope to include people across the USA. This afforded us the opportunity to interview interviewees about video calls *using* video calls. As the rest of the paper shows, autistic adults face a number of challenges when using video calls. We offered every interviewee accommodations to increase their comfort, including taking breaks during the interview to ease stress and cognitive load, turning off the video, and not calling attention to any neurotypical masking or lack of masking they did during the call. Some of the interviewees told us they felt more comfortable because we were explicitly asking them about their experiences with VC and autism, which is a subject that they were passionate to speak about with us.

The demographic distribution of our interviewees reflected some differences to the broader autistic population. 41% of our interviewees were women, which is higher than the 24% ratio seen in the wider US autistic population [31].[4] Though the wider autistic adult population has an 80% unemployment rate [4, 14], only 41% of our (non-student) sample is unemployed. In a study using qualitative methods, our goals are not to accurately represent the underlying distribution, but to uncover the broad range of diversity of perceptions, feelings, and experiences found in the population. We sample until we reach theoretical saturation, upon which we discover data that repeats themes. Our analysis identified a wide range of VC experiences with people that share many common autistic traits. As autistic people comprise a diverse spectrum, whenever possible, we report relevant autistic traits when they relate to our findings.

We developed a neurodiversity-sensitive model of computer-mediated communication, but with qualitative data are unable to quantify how each stressor and coping strategy affects the interviewees' preference for and effective use of each CMC channel. In addition, we believe that researching CMC from the lens of autism surfaces user needs that can situationally apply to neurotypical people as well. Thus, while we believe that these stressors and coping strategies may apply to a wide swatch of neurotypical video callers, further research is necessary to validate our model with them.

## 4 RESULTS

Here, we present the ways in which our interviewees currently use VC. All of our interviewees described ways that VC caused stress in the form of sensory sensitivities, cognitive load, and anxiety. They employed various coping strategies to manage those stressors, which we present within the structure of the flow of a VC: preparation, initiation, participation, and termination.

### 4.1 Video Calling Experiences

First, we report on current usage scenarios and strategies of adults with autism for managing social interactions over VC. We then describe ways in which autistic adults do hidden work during VCs to adapt to neurotypical social norms.

[4]The measured ratio of autistic women to men has risen over time as diagnostic processes improve.

*4.1.1 Usage Scenarios.* In general, our interviewees considered the term *video calling* to mean any technology used to speak and collaborate with others while one or more people are remote. Based on our interviewees' experiences, "remoteness" could be as close as down the office hallway, or as far away as another continent. All of our interviewees had used a variety of VC applications more than once to communicate with different groups of people in their lives. Primarily, they used VC applications such as Skype, Zoom, and WebX for conducting work or collaborating on school projects. In these settings, the interviewees' most often played the role of participant, although some interviewees had job roles in which they were responsible for hosting VC meetings for critical team functions (e.g., managing a daily team meeting to triage work items). Our interviewees used VC on desktops, laptops, and mobile devices. In some of these situations, interviewees chose to meet with their colleagues via VC even though they were geographically close to them. For instance, some interviewees used VC to "hang out" with other students while they did homework, choosing to do so over an extended VC rather than meeting in person. With regularly occurring meetings, such as daily work sync meetings, some interviewees occasionally chose to attend the meeting remotely, citing reasons like wanting to multi-task. Interviewees found VC to be preferable to FtF for social and logistical reasons, such as finding VC to be easier when both people were busy at night (P27), helping make lulls in conversation less awkward (P20), enabling the ability to control which part of one's face is shown on the video (P22), and maintaining personal space so no one gets offended (P06).

Interviewees valued having the option of participating in a VC while they were on the go on their mobile device. However, for long conversations or to collaborate for work, they preferred the stability and reliability of their desktops or laptops. Some also reported that they did not like the video streaming experience on mobile devices because their conversation partner would move around; this movement caused the video stream to show jarringly odd angles of their video partner's face or environment.

The majority of our interviewees also used VC for purposes unrelated to work and school, mainly for connecting with family, socializing with friends, and for entertainment. Several interviewees enthusiastically described participating in fan fiction and gaming campaigns on Discord[5]. Some interviewees enjoyed watching people stream themselves engaged in an activity such as drawing, on platforms such as Twitch.[6] In these scenarios, they were part of a large audience who was watching the performance. Only the performer was live streaming audio and video; the audience would communicate with each other and the performer via a live chat window. Finally, one interviewee used VC as a step in getting to know potential dating partners they met through Snapchat.

### 4.2 Stressors

In this section, we provide an overview of the stressors the interviewees experienced during VCs. Following this section, we provide more details about how these stressors manifest during the VC phases and the actions interviewees took to manage these stressors.

*4.2.1 Sensory Sensitivities.* Interviewees were keenly aware of managing the sensory inputs in their own physical environment, in addition to the streaming audio, video, shared desktops, chats, and the interface of the VC application. Streaming video elicited the strongest reaction by interviewees. Its availability was viewed by some interviewees as a way to *"have some of the advantages of in person conversations without actually being in person"* (P17). From the perspective of a receiver in a VC, some interviewees valued receiving nonverbal cues like gestures, facial expressions, and nodding, or even just seeing their conversation partner's face as it *"gave me something to focus on,*

[5]https://discordapp.com
[6]https://www.twitch.tv

*even though it is awkward."* (P21). Focusing on their partner's face provided stimulus, making it *"a little bit easier to focus if I am having trouble that day."* (P17).

However, some interviewees found watching their partner's face to be quite overwhelming, similar to how they felt in FtF interactions. In fact, they revealed that they tried to *fake* eye contact by looking in between their partner's eyes, on their forehead, or in the general direction of their mouth and nose. P11 offered:

> *"I've had trouble with eye contact. What I used to do was focus on people's mouths because to me it was easier, and it made more sense to begin with because the mouth is the part that's moving. Over the years I've gotten better at it, and now I just focus on the center of their face." (P11)*

Interviewees commented on the fact that eye contact in a VC is never aligned, a situation that they actually appreciated:

> *"I probably make eye contact too much. But when you're video conferencing nobody knows where to put their eyes anyways. So everybody doesn't make eye contact. It's great. If I'm staring [or] needing to look off into space, it's not as apparent." (P06)*

Interviewees were often disturbed by other sensory signals transmitted by the video, such as bright lights. Fluorescent and undimmable lights in their physical environment, especially in spaces they could not fully control (e.g., their work office), were bothersome. Notably, a few interviewees said they experienced regular migraines due to bright lighting or the blue light emitted by computer monitors. They could also be bothered by the amount of light transmitted by their conversation partner's camera view especially in contrast to their own dimly lit environment.

Besides video, the interviewees also mentioned concerns about the audio in VC. Because they listened to their conversation partner's voice for changes in tone, volume, and intonation as cues to the meaning of what they said, they found background noises, such as cars, typing, or eating noises, to be distracting. These distractions often caused them to inadvertently switch their focus to the noise and lose track of the conversation.

In VCs with multiple people, interviewees noted that they became overwhelmed when people talked over each other or repeatedly interrupted each other. In some cases, interviewees responded by matching their conversation partners' styles just to get their points across. P08 described that he may not realize he was interrupting others, saying that *"it is entirely possible that I interrupt sometimes and don't notice, and it never comes to my attention that I just did that."*

*4.2.2 Cognitive Load.* Interviewees described many activities and aspects of interactions that required cognitive and emotional processing at a level that impacted their ability to maintain social interactions. These included topic familiarity, distractors, developing a mental model of their conversation partner, managing the conversation, emoting to others, reading other people's emotions, and reading other people's body language. Their description of managing the cognitive demands of VC illustrate the amount of effort involved.

Interviewees discussed how they cognitively processed input channels, often describing that they did best by focusing on one channel at a time. For example:

> *"Audio-wise, I can track what you're saying and going like that but as soon as I have to factor in body language, I can either pay attention to your body language or I can pay attention to what you're saying." (P03)*

Interviewees discussed feeling overwhelmed with the impromptu nature of VC, especially for work teams that keep their VC application running in the background. Their preference was to use email or talk FtF if the issue was urgent. P08 shared that he is *"notorious among my colleagues for never running [VC] in the background. You have to organize with me ahead of time for [VC]"* (P08). A

common reason among our interviewees for not keeping a VC application running is that they found it difficult to immediately switch to the topic of the impromptu VC or other CMC, such as texts and application notifications.

Cognitive processing appeared to remain an important part of their experiences, even after they felt they had acquired new socio-emotional skills. P14 described this learning experience as follows:

> *"Part of the autistic experience is learning ... basically most social interaction at work, there's a certain amount of manipulation to it, and pattern analysis. Like we actually have to consciously think about it a lot. But if you do that long enough, you start memorizing the patterns, and you find little shortcuts, so over time it gets a little bit easier for some people." (P14)*

In summary, a VC required meta-cognitive and low-level cognitive processing about the content of the meeting, technology-mediated interruptions from other people, and the surrounding social-emotional environment.

*4.2.3 Anxiety.* A consistent theme across all the interviewees' VC and FtF experiences was managing anxiety. They described many contributing factors to anxiety, including their role in the conversation, social familiarity, topic familiarity, conversational goals, adhering to social norms, and their current socio-emotional capacity. At a topic level, they expressed that they were more comfortable talking about concrete, familiar "things," such as games, technology, or work deliverables, more-so than when talking about "people." Situations that required negotiating, conveying nuance, or being unsure created stressful emotions. Interviewees discussed how their emotions, especially stress, in technology-mediated social interactions correlated with their relationships and the goal of the interaction. In a closer relationship, their stress was lowered. However, if the goal of the interaction was related to a conflict or ambiguous task, stress rose, even if they were socially close to the conversation partner. Our interviewees conveyed that their past emotionally-laden CMC conversations included: planning a vacation with others, collaborating and presenting in a design meeting, interviewing for a job, and talking to someone after they had a baby. Note that these are situations that can heighten emotions for even neurotypical people.

Anxiety drove some interviewees to staunchly *avoid* using video web cameras, two said they did not even own a camera. These interviewees described feeling self-conscious about being on camera and felt judged by others. They felt less freedom to move around, multi-task, stim, and fidget. In addition, the experience of seeing themselves in the video camera was uncomfortable. P15 described it *"like staring at a giant mirror for an hour."*

During VCs, the interviewees strove to adhere to neurotypical social norms, such as eye contact, as described in the section above about Sensory Sensitivities. Since the way they expressed themselves verbally and through body language sometimes differed from those social norms, they had anxieties about being misunderstood, as P19 described:

> *"Someone may think that I'm either seeming upset if I'm thinking about something, an autistic person may feel like that. Or if I'm trying to say something in a different way, sometimes my voice tone may sound disrespectful when it meant to be respectful." (P19)*

They also worried about misreading their conversation partner because they would miss nonverbal cues such as *"the tightening of the face, a straining of the smile. Things like those I may not see unless I am actively looking for them."* (P19)

Interviewees described feeling significantly less anxious when talking to other autistic people. They felt more comfortable to be themselves because the social norms were more intuitive and they did not have to mask their autism. They could bond over common expectations about their

experiences. P03 found it helpful to talk with other autistic adults, saying *"I always get really curious about what their experiences are. Do they have the same? Do they not like this thing?"* (P03)

In summary, interviewees experience stressors in the areas of sensory sensitivities, cognitive load, and anxiety. This heightens their anticipation of a VC and their need to employ coping strategies, as we report on next.

### 4.3 Coping Strategies across Video Calling Phases

This section describes how, in response to the stressors previously described, interviewees developed strategies to use during a VC to support their sensory, cognitive, and social needs. These strategies consisted of physical actions and cognitive tasks that they performed before, during, or after a VC. Without appropriate coping strategies, interviewees reported becoming more stressed, less able to interpret social-emotional cues, and less effective in their role for the meeting. They reported investing a lot of effort in using such coping strategies.

*4.3.1 Coping Strategies during Preparation Phase.* Prior to VC, autistic interviewees make sensory, cognitive, and social adjustments to control for anticipated comfort levels and social expectations.

We saw a high degree of sensory awareness and respective adjustments for the purpose of self-sensory management, as well as a high degree of consideration for the sensory experience of other VC group participants. This includes adjustments to minimize physical discomfort, sometimes viewed as pain-reduction, and to improve sustained attention for themselves and others. For example:

> *"I try and find a comfortable place to sit. Like right now I'm sitting against the back of my bed so I have back support, and then I have a pillow on my lap so that I'm not totally bending over." (P15)*

To improve the sensory experience for themselves and other participants, both audio and visual adjustments were made. A concerted effort to perform microphone and speaker testing, in addition to making the resulting audio settings adjustments prior to the call, helped the autistic individuals control for overly loud interviewees, and avoid the social embarrassment and guilt of being disruptively loud for others.

Visually, many of our autistic interviewees would brighten or dim the lighting, including in-room background lighting and on-screen adjustments. Some interviewees mentioned blue light was especially painful for them, so they used the Windows 10 Night Light setting or third-party software to modify the color temperature of their screen in addition to dimming it. However, for others, blue light filtering was not enough, and they wished they could make more effective filtering and dimming adjustments:

> *" ... now it's night shift mode. But that only removes the blue ... what happens when you're still trying to read text? Now you just have a bright red blurt in your face ... you're limited to what the hardware is designed to do." (P16)*

To reduce cognitive overhead during the call, almost all autistic interviewees placed emphasis on technical reliability and general testing of the VC software well in advance. Preparing a medium for note-taking was one method of ensuring in-call content would be adequately processed and later recalled; this also seemed to help with focus. Personal comfort and physical and neurological calm were important and well-prepared for, also. Methods of pre-call anxiety-reduction include: hydration, background music, seating supports, position adjustment, and going to the bathroom in advance.

In VC, autistic people tend to refine their behavior and environment to meet anticipated social norms and expectations of others. Autistic meeting leaders, conversely, preferred to control social

dynamics by constructing a set of social meeting rules to be agreed upon by participants in advance. One example of such a rule was strict time-boxing, where participants who ran over an appropriate contribution window were muted by the autistic organizer:

> *"During the meetings, we allot each person a certain amount of time to speak. We let them know, and if we try to tell the person to stop and they won't, we mute them. It's a bit over-controlling but the members understand in the end that we have a limited amount of time for chapter meetings. Other people need to be heard, and it is not all about them, and especially if they may go off on tangents maybe they want to talk about school or something but it's not relevant. You can mute them individually." (P19)*

Some autistic people admitted to extensively researching future VC group participants online to help them formulate a mental model of other VC participants, leading to greater prediction of actions and confidence approaching those members during the call. Interviewees liked to turn their camera and microphone off before the start of the meeting, however acknowledging that these adjustments can cause increased anxiety when it comes time to contribute: *"I scramble to turn the microphone back on."* (P13) Interviewees were careful to hide personal belongings from view, and an emphasis on personal appearance and hygiene increased with formality of the call. Some interviewees who were more familiar with VC experienced less anxiety, and therefore made fewer preparations, especially with personal calls.

*4.3.2 Coping Strategies during Initiation Phase.* As the VC begins, there are continued sources of sensory, cognitive, and social concerns for our interviewees, all of which need to be managed. In general, they feel anxious about the initial *"two–three minutes of ... chaotic mess"* (P09), which includes disorganization as other VC participants join, and the pressure to adhere to awkward social pleasantries, namely, small talk. Interviewees expressed feeling nervous anticipation, and sometimes irritation, about audio, video, and connectivity issues with VC. There seems to be an exception for VC with close friends, in which technical issues are less of a concern. When entering a VC, many interviewees felt the immediate spotlight was uncomfortable. Here, P15 explains the contrast between in-person conversation and VC in regard to immediacy:

> *"[In person,] I won't know who's in a room [before I enter] ... but once I walk into the room, then I see people before I actually have to walk up to them and talk to them, you know? So at least I have a couple seconds warning. Whereas with video chat, boom I'm there." (P15)*

The majority of sensory concerns during the initiation phase are visual distractions and unhelpful pixelated visuals. P03 experiences the following camera anxiety:

> *"The new [VC] thing that I don't like is when you launch it, I always have ... anxiety on whether it's going to turn on my camera or not. Sometimes it does turn on the camera/turn off the camera. The first couple times it launched and turned on a camera, I was like, oh, this is no good." (P03)*

Some of this was controlled for by disabling the video prior to connecting. However, interviewees noted this can later cause confusion when it comes time to present their desktop. We saw similar feelings around microphone disabling and enabling during initiation.

Presenting is especially burdensome due to potential technical hiccups, such as poor bandwidth or a weak or lagging wireless connection. Such technical concerns consume cognitive resources at the time of VC initiation. General questions at the outset of a call include *"Are we on," "Can you hear me,"* and *"Can you see me"*. (P03) For this reason, the *"beginning is especially awkward."* (P03) These are sentiments that can likely be shared by all VC users; however, the effects appear to be amplified for the autistic segment. P07 reported stress for hours prior to a call, and fusses with her camera

for 30 minutes before initiation, to reduce awkwardness and mitigate technical complication. P15 also explains the discomfort of initiating VC:

> *" ... it's really awkward for the first five or ten minutes or so. And then, usually it becomes better as I sort of get used to the- I don't know how to say it. Just, I guess get acclimated maybe to the video call and all of that." (P15)*

Social coping during VC initiation is a natural continuation of the coping strategies employed in VC preparation: autistic adults often strive toward mental model completeness of their VC participant(s). In the initiation phase of VC, they do this via the avenues previously mentioned: small-talk, introductions, and role-establishment, much of which is an added source of pressure and anxiety. An interrupted VC initiation causes anxiety by reducing the autistic adult's sense of control and certainty, a natural side-effect of the then hampered predictability of their interaction: *"[missing introductions] puts me a disadvantage"* (P13). By creating a mental rendition of their conversation partner(s), autistic adults in VC can better calculate the optimal interaction style for each participant, which is consistent with the hyper-systematizing previously mentioned.

Autistic users are highly considerate and worried about the message being portrayed by their facial expression. Many are careful to plaster a smile on their face out of concern that their natural facial expressions may mislead the receiver into thinking that they are disliked. Some interviewees found the video preview very helpful because it showed them exactly what their conversation partner was seeing. P17 used her video preview to monitor her facial expressions. Many interviewees discussed that the misalignment of the camera views between them and their partner allowed them to *"fake eye contact by looking at my web camera, so the person thinks I'm looking at them, but I don't have to feel them looking at me"*. (P14) Note that these attempts to avoid discomfort, and the sending of incorrect conversational signals, is a highly conscious effort.

*4.3.3 Coping Strategies during Participation Phase.* After the initiation phase, the VC discussion shifts into the body of the meeting or the main topic or activity of a social VC. During the participation phase, the interviewees needed to actively moderate sensory inputs. Interviewees who perceive a lot of information from the tone of someone's voice desire a clear and consistent voice channel. They actively manage audio in the VC system to adjust speaker volumes—individually for each speaker if that was supported in their VC system. They adjusted the light coming from their partner's web cam to minimize painful bright lights. They engaged in repetitive behaviors like fidgeting, stimming, and walking around, which helped them release energy and focus on the content of the meeting. If they were running a meeting with people they knew well, they advised them about how to minimize sensory distractions. For example, for a monthly meeting, P19 told remote VC interviewees to stay in a quiet place, since otherwise background noise *"disrupts the entire meeting, and so we'd have to mute them"*.

Interviewees had a specific set of coping strategies to manage their attention during a VC. To focus on the content of the conversation, many interviewees turned off their own camera and the streaming video of their partner. By reducing their visual inputs, especially that of another person, they could better concentrate, as described by P09, *"if I'm trying to concentrate, then, I stare at my phone, and I can really absorb what people are saying, or I take notes at the same time."* A benefit of turning off their own camera was that they could multi-task and felt more comfortable engaging in repetitive behaviors. Interviewees could become distracted by the video stream of their partner when there were background movements or actions of the meeting participants, especially if they were doing something repetitively like typing. If a conversation partner was sharing their desktop, their mouse cursor movement could be distracting. Some interviewees stressed that they were audio learners and relied on their auditory skills to read the tone of the speaker's voice, and therefore, felt they were more effective communication partners with the video stream off.

Interviewees were actively engaged in meta-cognition tasks during the participation phase. Several interviewees described ways that they externalized knowledge to help them process the content of the meeting. Their strategies included taking notes, looking at off-line copies of presentation slides, or reading the VC chat window—rather than focusing on the verbal discussion. Some interviewees had difficulty tracking topics, sometimes because they were still thinking through a point of a previous topic or formulating their thoughts because they wanted to express them thoroughly. Tracking topics was also difficult due to notifications from other applications or the VC system, which caused them to context switch. In those cases where the VC connected with a conference room with multiple participants, our interviewees experienced stress and cognitive load trying to identify who was talking.

There were some scenarios in which interviewees were concentrating on one topic and did not seamlessly progress to the next topic of discussion. For example, if they were discussing a topic that they were very interested in, they could become so engrossed that they could speak at length. Due to their excitement, they sometimes did not pick up on their partner's nonverbal cues about wanting to chime in or change the subject. In the scenario of listening to a VC, interviewees may be contemplating a point from one slide, not noticing that the presenter has moved onto the next slide.

In terms of coping strategies to support social relationships, interviewees were conscious of how they were being perceived by others. On one hand, a few interviewees expressed that they like how they come across in VC, such as P03 who said, *"virtual me is better and more dynamic than in-person me."* However, the majority of interviewees worried that they would be misunderstood or harshly judged over VC. The VC video preview was a useful tool to check in on their facial expressions and their body position.

> *"You can kind of know exactly what the person is seeing as well which is sort of unique because I can see, there's a preview at the bottom where I know exactly what's going across to the other person and that's very helpful for me." (P17)*

A surprising source of social anxiety was the usability difficulties of the VC application beyond technical connectivity issues. One interviewee described at length how anxious she got sharing her desktop, knowing her computer actions were being observed. She worried that her teammates would judge her for experiencing usability issues such as accidentally clicking on the wrong button or being unable to locate the "stop sharing" button:

> *"It is uncomfortable using my desktop incorrectly while I'm sharing desktop. I always forget where the share button is... Watching someone else do it, it seems so easy." (P13)*

Interestingly, interviewees had different perspectives on what constituted the efficient use of time during VCs, mostly depending on the purpose of the VC. For work related VC, interviewees generally perceived that VC conversations were more concise than they are in FtF situations in which people can talk for unpredictable amounts of time and meander into unanticipated topics. On the other hand, interviewees used VCs to hang out with people they are close with and with fellow students for online study dates. In these situations, they appreciated being able to take pauses and take advantage of being in their own homes to take care of other tasks or lie down on the bed to relax before resuming the VC.

*4.3.4 Coping Strategies during Termination Phase.* Closing out a VC presents yet further challenges for autistic interviewees. In terms of sensory sensitivities, the main issue reported in this phase was that interviewees would "run out of their spoons" before the end of the actual conversation. This can lead to their abruptly leaving the call. This cognitive load exerted in this phase also extends past the end of the actual VC. Interviewees reported being anxious about ambiguous action items

and, when the call involved new people, what their next interaction should involve. However, a positive aspect of VC is that they can refer back to the associated chat window to help remember what was said and to find notes and links for follow-up action.

The same awkwardness and uncertainty about proper social etiquette that are evoked in the previous phases persist and can even be heightened by the fact that once the VC is over, no further time is available to correct or adapt. P17 described this uncertainty as follows:

> *"'Cause it's different than a conversation. And I've noticed that in the video calls that I've been in recently at least, I always feel like I end it slightly sooner than I'm supposed to 'cause I'm like okay thank you, bye. And then I click end and I think maybe I'm supposed to wait longer, I don't know. That's a weird etiquette part that I feel like I might be doing wrong." (P17)*

### 4.4 Masking Autism, Compared to Interactions Among Autistic People

Across all VC phases, the most persistent, and consistently reported, social coping strategy was trying to mask being autistic by attempting to adopt neurotypical behaviors and expressions. In addition to "masking', interviewees used phrases such as *"faking NT [neurotypical]"* (P03) and *"passing"* (P01) for this behavior. Masking actions were verbal and nonverbal, including responding to small talk, laughing at a joke they did not understand, maintaining eye contact, and hiding verbal tics. Masking behaviors required high cognitive load and were *"exhausting"* (P03).

Interviewees explained that they did not feel the need to mask when they were with (1) close family and friends, (2) people they felt would accept their autistic behaviors, and (3) other autistic people. As P09 stated, *"it's easier for me to talk to people on the spectrum than those who aren't because they're less likely to take things personally. It's more face value with them."* Some interviewees felt *"immediately very, much more comfortable"* (P17) and less self-conscious with other autistic people because they shared common experiences and behaviors. In a VC, P16 noted that *"there's not this expectation for you to look at the screen or do certain things because they're probably going through their own things."*

The contrast between the behaviors and feelings when autistic adults are masking, as compared to being with other autistic people, gives us insight into the natural ways of interacting that are more comfortable for autistic individuals. Our findings surface the social-emotional work autistic adults do to traverse between neurotypical and neurodiverse social environments.

## 5 DISCUSSION

In answer to our first research question (RQ1), we found that a primary benefit of VC for our interviewees was being able to engage in work, education, and social activities from the comfort of their own home. Their personal space is familiar and predictable, making them feel better situated to face the unknowns of online classes, social events, etc. They found VC to be less anxiety-inducing than being FtF because it was more contained. Interviewees also appreciated that the VC interface provided some structure to the interactions. For example, some VC interfaces display the video of the current speaker larger than other people's videos, which made it easier for our interviewees to focus on the speaker and what they are saying.

Our interviewees' experiences with video calls showed us that CMC affords them opportunities to engage in social interactions across a variety of CMC channels. In some ways, communication with others is even enhanced over FtF communication because it affords autistic people the ability to use the technology to ease their anxiety, reduce their cognitive load, and better manage their sensory sensitivity. We introduce our neurodiversity-sensitive computer-mediated communication (NDS-CMC) model to pull all of these factors into a single conceptual process (see Figure 1) that

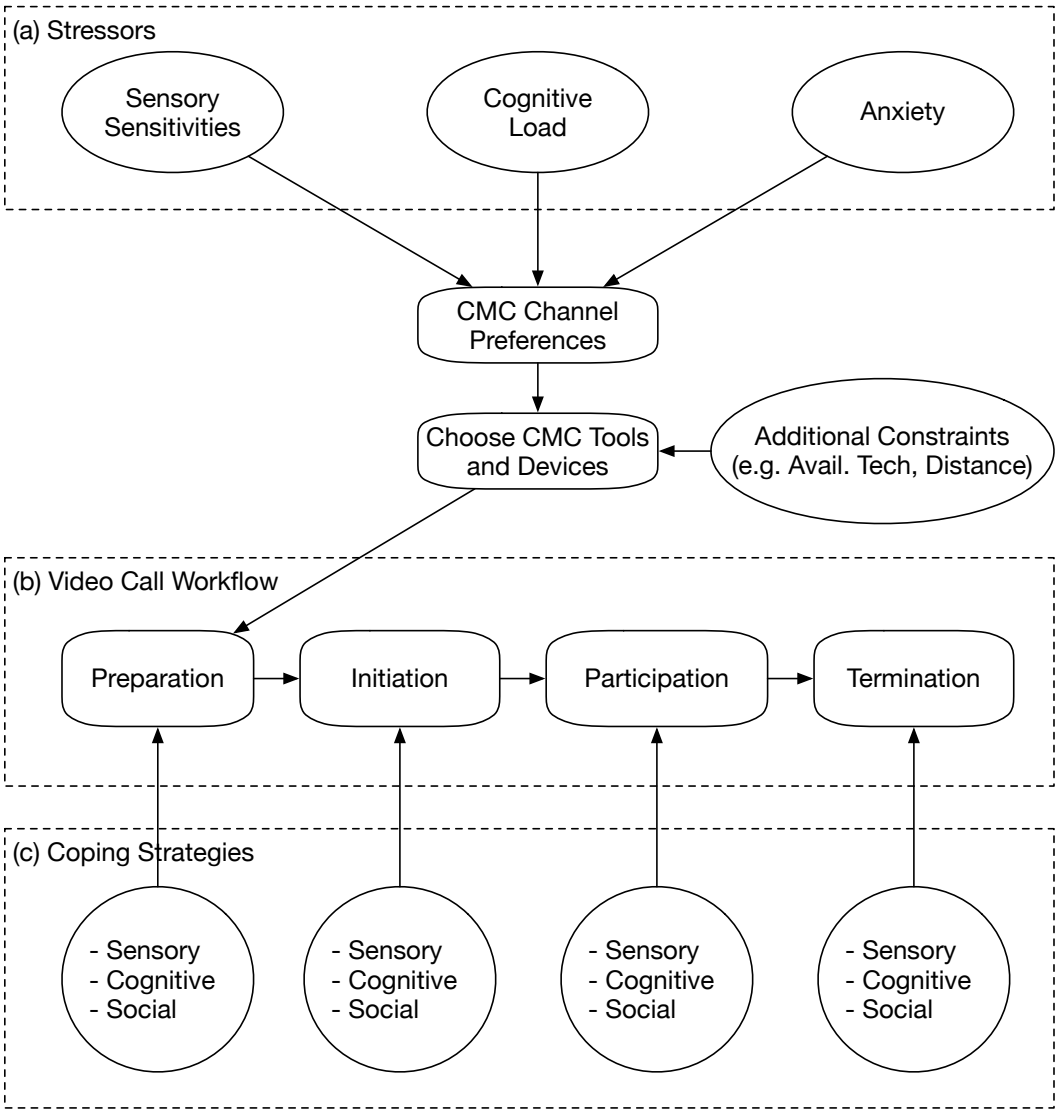


Fig. 1. A neurodiversity-sensitive computer-mediated communication model (NDS-CMC): (a) Anxiety, cognitive load, and sensory sensitivity are stressors that affect autistic users. These factors influence which CMC channels are used in a video call, in additional to the usual constraints. Once decided, autistic users invest in a variety of (c) sensory, cognitive, and social coping strategies to handle each of the (b) 4 stages of the video call workflow.

can help us understand their effects. We then expand on several of the more interesting findings and relate them to theory and prior literature to help explain their presence and their impact on CMC. Finally, we discuss how conversations among people on the autism spectrum mediate factors that affect the model.

### 5.1 Neurodiversity-Sensitive Model of Computer-Mediated Communication

We summarize our findings around our other research questions in Figure 1, which we call the neurodiversity-sensitive computer-mediated communication model (NDS-CMC). The NDS-CMC model depicts critical factors in (a) CMC selection and (b) VC workflow process for autistic individuals. The model shows that CMC channel selection is based on (a) key stressors—anxiety, cognitive load, and sensory sensitivity. In addition, their selection is influenced by constraints raised by prior CMC selection models, such as availability of technology due to socio-economic factors and physical distance to other participants. Once a CMC is selected, autistic users invest in a variety of (c) sensory, cognitive, and social coping strategies to handle each of the (b) 4 stages of the video call workflow.

The NDS-CMC model builds upon the theoretical foundations of Rich Media theory and the social theories of CMC (e.g., Social Information Processing and hyper-personalized model), yet differs in several important ways in the core areas of channel affordances, interaction partner, and interaction topic. First, the model contributes the notion of stressors to channel selection based upon our findings that users assess a channel for their ability to mediate stressors while using it. In essence, a channel's affordances include the user's ability to mediate stressors, such as ability to engage in repetitive behaviors. Second, the NDS-CMC model deepens our understanding of the importance of the social connections between the interaction partners in work and personal settings. Prior models discuss the ways that social ties can deepened or be hindered by CMC. Our model uses a broader lens of social norms, incorporating the impact of a participant's stressors and coping strategies, as we describe below. Last, the NDS-CMC model describes ways that the interaction topic impacts a participant's stressors, which, to our knowledge, is an area previously unexplored in prominent CMC theories.

As we examined the data from the interviews we conducted, we noticed three recurring sets of stressors that generate discomfort during VC (RQ2): anxiety, cognitive load, and sensitivities to sensory stimulation. Interviewees identified how these stressors influence their preferences for the CMC channel(s), tools, and devices they used to speak with family, friends, and work colleagues (RQ3). After committing to a CMC channel for a call, the interviewees spent considerable effort coping with the consequences of their CMC choices (RQ4). They employed a variety of strategies to ensure a more comfortable sensory, cognitive, and social context in each of the four stages of a video call: preparation, initiation, participation, and termination.

Most interviewees mentioned some degree of sensory sensitivity to light or sound caused by their own environment or by the environment around their conversation partners. In addition, they frequently had to manage their own excess energy through fidgeting or stimming.

Factors that affected anxiety included familiarity with the conversation partner(s), familiarity with the topic, the social norms required for the conversation (e.g. were pleasantries and small talk expected?) and whether they had enough socio-emotional and cognitive resources to uphold those norms (i.e., minimizing autistic behaviors through neurotypical masking). Other factors include their role in the conversation and expectations about their goals, and whether the conversation was about a person involved in the conversation (especially if it was about them) or whether it was about a non-personal topic.

Finally, cognitive load could also be affected by familiarity with the topic of conversation, their abilities to process the content (e.g., visual aids, fonts, verbal, and written supporting materials), the completeness of their mental model of their conversation partner, their abilities to look their conversation partners in the eyes, read their emotions, or emote their own feelings during the conversation (interviewees reported that each of these takes deliberate cognitive effort), their speed in deciphering ambiguous conversational content (e.g., which Thursday does "this Thursday" mean

if today is Friday?), the overall efficiency of the conversation (i.e., how quickly do people make their points?), and very importantly, their abilities to resist paying attention to visual and audio distractions during the call.

These factors weigh heavily for many of those on the spectrum when considering whether to include a text, audio, or video channel in an upcoming call with someone else. Typically, to enjoy a more comfortable conversation, their goal was to minimize anxiety, reduce cognitive load, and to effectively manage sensory overload. For example, in less-stressful situations, interviewees were much more likely to choose to want to interact with the maximum bandwidth possible (which was often FtF) but would choose video calls when distance and technology constrained them. In situations where expediency of conversation was at the forefront (again with low stress), texting or email channels would be preferred. When someone was less well known, choosing a video call heightened an autistic person's anxiety, increasing their preference to turn off the video channel, or even choosing texting or email to help reduce it. Various factors' impact on CMC channel preferences will be discussed in the next section.

This model helps to explain the CMC preferences of those who place themselves somewhere on the neurodiversity continuum. Conversations, however, are co-constructed from many people and their preferences. The social and cognitive norms which are co-constructed and negotiated between all parties in a conversation were not brought up by many of the interviewees. However, they were reflected in the interviewees' own anxieties and cognitive difficulties with video calls because most of the recounted conversations lacked any specific communication required to establish norms, thus putting them at the mercy of all parties' unvoiced conversational expectations.

CMC channel selection is also influenced by typical factors, for example constraints on available technology (e.g., does a person have access to Skype or can they find a place to sit for the meeting?) and opportunities for meeting conversation partners face to face. These factors can trump the neurodiverse factors we have introduced in the NDS-CMC model because they more directly affect the capability of the conversation to employ particular CMC channels.

Next, we discuss some of the stressors in more detail.

### 5.2 Impact of Stressors on Computer-Mediated Communication Preferences

Our results showed numerous examples of how stressors could help and hinder the comfort level and effectiveness of autistic adults engaging in video calls. Though an important goal of VC is reducing barriers to increase closeness (and thus, propinquity [30]), our autistic interviewees identified many positives to maintaining distance. Informing RQ3, we found that the technical affordances of CMC enable autistic people to have more control over their sensory environment, relieve some of the anxiety caused by close contact with people they do not know very well, and reduce the cognitive load incurred as they process and follow the content of their CMC-facilitated conversations.

Interviewees preferred to switch to low-bandwidth CMC channels to help them manage sensory overload. If you cannot see the other person, you cannot be overwhelmed by the light of the bright window in that person's background. If you remove or mute the audio channel, you will avoid the distraction of the incessant hum of their computer fan which was positioned too close to their microphone. Sensory sensitivities are explained well by Weak Central Coherence theory: autistic people's hyperfocus on details applies equally to their perceptual system, overloading their senses with input that they cannot process. For them, it actually hurts to look at bright lights or listen to loud sounds, which would merely be an annoyance to neurotypical people. Low-bandwidth CMC channels also make it possible to stim and fidget as much as needed without having to worry that the other person can see it and will judge them for it. Similarly, when high cognitive load was caused by stressors, interviewees indicated that they would prefer low-bandwidth and

asynchronous CMC channels (e.g., text or email) to minimize the need to perform tasks that were cognitively-demanding for them. Scholl et al. found that people preferred text-based chat over audio because it was less intrusive, and was easier to use, especially for communicating in a second language, something that incurs more cognitive load than using a native language [45]. As autistic people "use up spoons," their socio-emotional cognitive resources decline, making tasks even more difficult and sometimes leading them to abandon the conversation. Incidentally, VC's inability to support direct eye contact is actually an *advantage* for most autistic people, who find it cognitively draining to do this when speaking to someone FtF.

Interviewees described a complex dynamic between their social-emotional skills and their agency in the conversation. In terms of expressing their own emotions, their emotional affect did not always come naturally; rather, they had to make a conscious effort to emote both positive and negative emotions. Their lack of facial expression was sometimes interpreted by their neurotypical conversation partners as anger, even though they were feeling happy. Since this negatively affected the conversation, autistic adults would spend considerable cognitive effort to deliberately make facial expressions just for the benefit of their neurotypical partner. Their capacity to emote was impacted by their sensory comfort levels, their energy levels, and their relationship with their conversation partners. One way to conserve energy was to rapidly move their eyes between several people in a conversation, minimizing the time spent looking at any one person's eyes but still believing that the others felt they were looking at them.

Hyper-systemizing theory helps explain the cognitive difficulties experienced by autistic adults when trying to keep up with the flow of a conversation, especially an inefficient one. The flood of (potentially irrelevant) details in a conversation requires intense cognitive effort to manage; their ability to process input in larger, more abstract chunks is limited. When trying to read others' social-emotional cues, our interviewees intentionally scanned for nonverbal cues—something that was difficult for them and which lowered their focus on the verbal conversation. Their lack of a theory of mind explains their difficulties in reading the emotions and body language of other people. However, we observed something more. Interviewees reported having an easier time reading positive emotions and body language (such as happiness and laughter) than recognizing anger or sadness. This may be because people tend to suppress negative emotion expression in social interaction [26, 50]. They said it was especially difficult to read emotional cues when the sender's words where incongruent with their tone of voice and body language. This occurs when someone is being sarcastic or hiding their true feelings.

Interviewees also told us of the need to mask their autistic behaviors to conform to neurotypical politeness norms, even though it required intense cognitive effort. Tannen showed that when conversing, neurotypical people try to match their *conversational style*, a concept that includes vocal pace, prosody, and relative volume [49]. For the autistic partner who has trouble perceiving these vocal characteristics, it would be difficult to please their partner by matching them. A similar challenge occurs with misinterpreted and mismatched facial expressions and body language. Several interviewees reported that their neurotypical conversation partners would often misinterpret their lack of facial expression as anger, even though they were feeling happy.

Interviewees told us that the stressors that induce anxiety were a constant worry for them, and influenced their CMC preferences. Only when interviewees were familiar with their conversation partners and the topics of conversation did they report feeling most comfortable in an in-person or high-bandwidth video call. Otherwise, they made use of low-bandwidth CMC channels like texting or email to keep conversation partners at a distance, relieving their anxiety over the impending situation. This fits with a theory of mind and the hyper-systemizing theory, which explain that when having difficulty anticipating the responses of a conversation partner, autistic adults will fear that saying the wrong thing will cause the conversation to break down.

### 5.3 Coping Strategies

Once autistic adults entered into video calls, they applied a variety of coping strategies to modulate their sensory, cognitive, and social needs to their desired level of comfort (RQ4). At each stage of the call—preparation, initiation, participation, and termination—interviewees revealed to us how their strategies interacted with their autistic traits, which we can understand with various theoretical constructs.

In the VC preparation stage, autistic adults spent considerable effort to manage sensory overload, dimming lights and managing ambient and distracting noises. Weak Central Coherence theory explains this desire to manage their ability to maintain focus on the conversation. Some interviewees also spent effort to sanitize the background visible on camera to conform to social expectations of the people they were calling. Finally, while persistent technical glitches are well-known to frequent users of VC, we believe our interviewees' extreme conscientiousness (one of the Big 5 Personality traits [33]) leads them to spend much more time than neurotypical people preparing the VC technology for meetings.

In the VC initiation stage, CMC channel selection was a major concern. Interviewees preferred low-bandwidth CMC channels because they felt they could do away with small talk and pleasantries, something that autistic people find difficult to do and which was also mentioned by Scholl et al. [45]. Almost every interviewee mentioned their need to control distractions by turning off their camera. This enabled them to reduce the anxiety and cognitive load caused by worries about self-presentation [27].

In the VC participation stage, interviewees again had to watch out for and explicitly manage sensory overload due to dynamic changes in sound and lighting in the audio and video channels of the people with whom they were speaking. In addition, the autistic person's propensity to repetitive motion explains their need to fidget and stim during the conversation to manage excess "energy." Cognitive load limitations and focus issues led many interviewees to explicitly externalize knowledge during the call; this helps them solidify their understanding of the conversation topics, according to the levels of processing effect [15]. Their difficulty following conversation topics relates to their attention to detail in a conversation (as is described by Weak Central Coherence theory); as they spend effort understanding what their conversation partners are talking about and obsess over choosing the right response, they fall behind and are unable to keep up as the topic of conversation changes. Finally, their need to fit in and fear of being judged for their autism led many to mask themselves to appear as neurotypical as possible [28].

Lastly, in the VC termination stage, interviewees spoke about "running out of spoons," forcing some to end conversations early in order to conserve precious socio-emotional cognitive resources. After calls finished, some interviewees involved in work meetings had difficulty clearly identifying action items; ambiguities were difficult to resolve because they found it too difficult to pay enough attention to listen to and remember the entire conversation while taking notes at the same time. The cognitive resources required to handle the details they notice (explained by Weak Central Coherence theory) exceed their capacity.

### 5.4 How Interactions Among Autistic Adults Impact the NDS-CMC Model

Our interview findings indicate that stressors and coping strategies in the model we have described can be mediated by the relationship between the conversation partners, especially when they are all on the autism spectrum. Our interviewees mentioned that they have a special skill in recognizing when their conversation partner might also be on the autism spectrum. Interviewees reported identifying autistic VC partners based on observing physical and vocal mannerisms in others that they regularly identify with themselves. Conversations with other people on the spectrum

significantly attenuated all three kinds of stressors, thus increasing interviewees' preferences for in-person or high-bandwidth CMC channels. For example, the need to mask autistic symptoms is unnecessary when the person you are speaking with understands all of your traits, does not judge you for them, and is able to understand what you are trying to convey without difficulty. Not only is there no need to incur the cognitive load of putting on a neurotypical mask, conversations could take place in the dark while pacing back and forth to minimize visual overload and manage excess energy without worry that the conversation partner would not understand. Information is conveyed purely by textual or verbal means, rather than being tacitly conveyed through emotion or body language.

In addition, some autistic interviewees told us that their ability to infer another autistic person's thought processes can be much more accurate because they are so similar to their own. Our interviewees told us that with neurotypical conversation partners, however, they must construct a unique mental model, adapted from a "generic model" for a person, that can anticipate their behaviors to various inputs. With those they interact with most often, they develop very complex sets of rules that eases the cognitive load required to predict their reactions while speaking with them. This challenges Baron-Cohen's proposition that most autistic people lack a theory of mind [7]. Perhaps, autistic adults develop adaptive skills to help them understand the thoughts of others.

## 6 PERCEPTIONS OF CMC AFFORDANCES AND DESIGN DIRECTIONS

Our final research question (RQ5) asks how VC tools might change to better meet the needs of autistic users. During our interviews, we asked the interviewees for their impressions of broad design directions and for any ideas they had for enhancing their VC experiences. Through those discussions and by analyzing the experiences of autistic adults using VC, we learned that their use is atypical in some important ways. However, these differences illuminate challenges that also affect most neurotypical people to a lesser degree and in specific contexts and across populations, such as potentially for non-native speakers. Below, we identify the key affordances of currently available CMC tools to highlight the value of these designs. Next, we describe design directions that emerged during the analysis of our research that could help improve the user experience for autistic users. Note that VC experiences are *co*-constructed among all of the VC participants. Therefore, it is important that all participants, autistic and neurotypical, work *together* to make the VC experience more comfortable for everyone.

### 6.1 CMC Affordances

Hogan posited that the use of CMC may prevent people from presenting their true self online [27]. Since the use of CMC affords autistic people the opportunity to choose the modality in which to interact, they are better situated to cope with stressors. In fact, they are *more* free to present their authentic selves to conversational partners. In addition, by minimizing the CMC bandwidth, they can maintain this authenticity over longer conversations because CMC enables them to better manage the need to mask their autism and conserve their socio-emotional cognitive resources.

CMC affordances supporting attention are critical for autistic adults. CMC applications could better help them focus on the important parts of a conversation (i.e. what their conversation partners are saying to them) by providing ways to filter or limit distractions (e.g. other conversations, moving objects, or interesting objects in the background). In dividing their attention between all of the distractors, it becomes impossible for them to pay enough attention to the conversation at hand. This impedes their abilities to follow conversations and cognitively process their content. Fortunately, the combination of CMC and other technologies can already help autistic adults in the filtering process. For example, Skype's ability to blur the background behind a person in a video call can relieve the need to pay attention to those background details. Attention pressure could also be

supported using lower-bandwidth CMC channels to relieve the need to maintain eye contact. As our interviewees told us, they would be able to shift their gaze around the room, and in doing so, increase their ability to focus on the conversation.

### 6.2 Making Social-Emotional Cues More Concrete

Our research highlighted the ways in which social-emotional cues generated by neurotypical people are complex and often ambiguous for autistic people, due to possible difficulties with a theory of mind. Our interviewees were optimistic about the design direction of technology to help people translate the social and emotional information that is being communicated to a form that is easier for them to see and understand. This direction could be supported by machine-learned classification of verbal and nonverbal cues, which could help CMC in a number of ways. First, in the case that someone finds visual information in a VC overwhelming and minimizes the window, covers it up, or otherwise avoids looking at it, computer algorithms could allow them access to simplified signals or summaries about the activity that is going on. In prior work, facial expressions were simplified into a bubble visualization to indicate what a person might be feeling [32], a technique that might benefit autistic users. Second, technology could provide a way for people to reflect on rich information after the fact, which could potentially be used as a training or teaching resource. For example, Boyd et al. created SayWAT, a tool that gave autistic users feedback about their vocal prosody in FtF conversations [10]. Washington et al. provided emotion-recognition training for autistic children using a Google Glass wearable device that could automatically recognize other people's emotions [55]. Benssassi et al. presented many ideas on adapting wearable assistive technologies for use by autistic users to help them read others' emotions [9]. Many of these training-focused technologies could be adapted to operate within the context of a VC. Third, we found that social-emotional cues are ripe for misinterpretation by our interviewees, especially when a person's words, voice tone, and body language are incongruous. Multi-modal classification could help identify incongruities in subtle cues and flag them. This could also be used to highlight moments that require closer attention.

Making emotional and social cues clearer could increase everyone's confidence and agency. However, our interviewees stressed that it is important that this information is not presented in a prescriptive manner. Rather, it should augment the individual's understanding of the situation and help them make decisions. For example, it can help them to decide when to transition from one topic to another, or when to end a meeting. Byun et al. created a system that used gestural and nonverbal cues to indicate to VC participants how well their conversations were going [13]. Algorithms may still miss subtle cues and/or misinterpret expressions; there is still a long way to go before machines are close to the level of a human at this task. Nevertheless, the output from automated coding can still be useful.

### 6.3 Providing Conversational Assistance

Another key theme in our research is that conversational dynamics, such as ensuring everyone is contributing and handling interruptions, are ambiguous and can induce conflict. Currently, the CMCs used by our interviewees simply transmit people's verbal and nonverbal conversational cues without translation. However, these cues require autistic adults to consciously expend cognitive effort to interpret them. Our interviewees responded positively to design directions that could help alleviate this effort. For example, artificial intelligence (AI) technology could be used to transform verbal, nonverbal, and textual conversation cues into an easier-to-read form. In essence, an AI could act as a co-facilitator of a meeting. For example, the meeting agenda could be added to a CMC interaction, and then transformed into real-time reminders and explicit notifications about the current topic. AI tools could make clearer indications of who is currently speaking, even when

there are multiple people co-located in a single conference room. The AI could mediate issues around monopolizing the conversation, interruptions, and cross-talk. There could be explicit cues to indicate a speaker queue, and when to transition to a new speaker. To assist with comprehension, AI technology could provide the meeting content in alternative modalities like live captioning. In addition, captions could be made available after the meeting, appended with additional content that was shared during the conversation. These real-time and post-interaction conversational tools could help the users to be able to focus more of their attention on the content of the meeting and relationship-building.

Finally, our design ideas have the potential to increase awareness among VC participants of the diverse cognitive and communication styles of all of their collaborators, autistic and neurotypical. As teams work together, they would be more likely to explicitly negotiate social norms that would be compatible with all of their members, and more socially inclusive [58]. As Burke et al. suggests, establishing training programs for workers that includes social etiquette, diversity, and best practices for conducting effective VCs could help teams to co-construct mutually beneficial team norms [12].

## 7 CONCLUSION

Our research illuminated the often-hidden effort of autistic adults as they engage in video calling. They actively engage in coping strategies during a VC and other CMC to manage their main stressors: anxiety, cognitive load, and sensory sensitivities. The NDS-CMC model describes how these stressors impact the choice of CMC channel for autistic adults and makes visible the coping strategies they employ to participate in CMC. This work contributes to a growing understanding of the lived experiences and socio-technical practices of adults with autism. In future work, we will explore the technical and socio-technical design directions from this research incorporating the feedback from our interviewees and insights from this research. CMC technologies other than VC likely also present socio-technical access barriers to autistic people. Investigating and remedying these issues are areas for future exploration.

Our research revealed insights that have broad applicability because using autism as a lens for studying the VC experiences enabled us to identify some user needs that are critical for autistic users and inherent in everyone. Because of the sensory sensitivities that autistic people experience, they foregrounded distractions that probably bother all VC users; although, they may not be aware of their impact. For example, stories about heightened awareness of distractions from background sounds (e.g., someone eating or typing) suggested audio filtering techniques that would be appreciated by all VC users. Similarly, comments about the visual distractions in the background of a video call would be addressed by background blurring features that are recently becoming available in VC, suggesting a more universal interest in that feature. These examples show how the heightened sensitivities of autistic adults can identify removing distractions that would be relieving to the more general user population. The goal of supporting neurodiverse VC means that the VC environment can be optimized for people with different cognitive styles, rather than a goal of making VC experiences match an idealized view of FtF interactions according to neurotypical expectations. In essence, improving the VC experience for autistic adults can make the VC experience more comfortable for all users.

## ACKNOWLEDGMENTS

We thank the participants of our study for sharing their experiences. We also thank the Accessibility User Research Collective (AURC) for assisting us with recruiting.

## REFERENCES


[1] Meryl Alper. 2018. Inclusive sensory ethnography: Studying new media and neurodiversity in everyday life. *New Media & Society* 20, 10 (Oct. 2018), 3560–3579. https://doi.org/10.1177/1461444818755394

[2] Hala Annabi, Karthika Sundaresan, and Annuska Zolyomi. 2017. It's Not Just About Attention to Details: Redefining the Talents Autistic Software Developers Bring to Software Development. In *Proceedings of the 50th Hawaii International Conference on System Sciences*. University of Hawaii at Manoa, Honolulu, Hawaii, 5501–5510. http://scholarspace.manoa.hawaii.edu/handle/10125/41827

[3] American Psychiatric Association. 2013. Diagnostic and Statistical Manual of Mental Disorders (5th ed.).

[4] Robert Austin and Gary Pisano. 2017. Neurodiversity as a Competitive Advantage. *Harvard Business Review* 95, 3 (May-June 2017), 96–103.

[5] J. Baio, L. Wiggins, DL Christensen, and et al. 2014. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR Surveill Summ 2018* 67, SS-6 (2014), 1–23. https://doi.org/10.15585/mmwr.ss6706a1

[6] Simon Baron-Cohen and Matthew K. Belmonte. 2005. Autism: a window onto the development of the social and the analytic brain. *Annu. Rev. Neurosci.* 28 (2005), 109–126. http://www.annualreviews.org/doi/abs/10.1146/annurev.neuro.27.070203.144137

[7] Simon Baron-Cohen, Alan M. Leslie, and Uta Frith. 1985. Does the autistic child have a "theory of mind"? *Cognition* 21, 1 (1985), 37–46. https://doi.org/10.1016/0010-0277(85)90022-8

[8] Roy Baumeister, Ellen Bratslavsky, Mark Muraven, and Dianne Tice. 1998. Ego Depletion: Is the Active Self a Limited Resource? *Journal of Personality and Social Psychology* 74, 5 (1998), 1252–1265.

[9] Esma Mansouri Benssassi, Juan-Carlos Gomez, LouAnne E. Boyd, Gillian R. Hayes, and Juan Ye. 2018. Wearable Assistive Technologies for Autism: Opportunities and Challenges. *IEEE Pervasive Computing* 17, 2 (April 2018), 11–21. https://doi.org/10.1109/MPRV.2018.022511239

[10] LouAnne E. Boyd, Alejandro Rangel, Helen Tomimbang, Andrea Conejo-Toledo, Kanika Patel, Monica Tentori, and Gillian R. Hayes. 2016. SayWAT: Augmenting Face-to-Face Conversations for Adults with Autism. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, Santa Clara, California, USA, 4872–4883. https://doi.org/10.1145/2858036.2858215

[11] Erin Bradner and Gloria Mark. 2002. Why Distance Matters: Effects on Cooperation, Persuasion and Deception. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. ACM Press, New York, NY, USA, 226–235. https://doi.org/10.1145/587078.587110

[12] Moira Burke, Robert Kraut, and Diane Williams. 2010. Social use of computer-mediated communication by adults on the autism spectrum. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10*. ACM Press, Savannah, Georgia, USA, 425. https://doi.org/10.1145/1718918.1718991

[13] Byungki Byun, Anurag Awasthi, Philip A. Chou, Ashish Kapoor, Bongshin Lee, and Mary Czerwinski. 2011. Honest signals in video conferencing. In *2011 IEEE International Conference on Multimedia and Expo*. IEEE, Barcelona, Spain, 1–6. https://doi.org/10.1109/ICME.2011.6011855

[14] June L. Chen, Geraldine Leader, Connie Sung, and Michael Leahy. 2015. Trends in Employment for Individuals with Autism Spectrum Disorder: a Review of the Research Literature. *Review Journal of Autism and Developmental Disorders* 2, 2 (June 2015), 115–127. https://doi.org/10.1007/s40489-014-0041-6

[15] Fergus Craik and Robert Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour* 11, 6 (1972), 671–684.

[16] Emmelyn A. J. Croes, Marjolijn L. Antheunis, Alexander P. Schouten, and Emiel J. Krahmer. 2019. Social attraction in video-mediated communication: The role of nonverbal affiliative behavior. *Journal of Social and Personal Relationships* 36, 4 (April 2019), 1210–1232. https://doi.org/10.1177/0265407518757382

[17] Emmelyn A. J. Croes, Marjolijn L. Antheunis, Alexander P. Schouten, Emiel J. Krahmer, and Daniëlle N. M. Bleize. 2018. The effect of interaction topic and social ties on media choice and the role of four underlying mechanisms. *Communications* 43, 1 (March 2018), 47–73. https://doi.org/10.1515/commun-2017-0046

[18] Richard L. Daft and Robert H. Lengel. 1986. Organizational Information Requirements, Media Richness and Structural Design. *Management Science* 32, 5 (May 1986), 554–571. https://doi.org/10.1287/mnsc.32.5.554

[19] Joyce Davidson and Victoria L. Henderson. 2010. 'Travel in parallel with us for a while': sensory geographies of autism. *Canadian Geographer / Le Geographe canadien* 54, 4 (Dec. 2010), 462–475. https://doi.org/10.1111/j.1541-0064.2010.00309.x

[20] Paul Ekman (Ed.). 2015. *Emotions in the Human Face* (2 ed.). Malor Books.

[21] Kathleen E. Finn, Abigail J. Sellen, and Slyvia B. Wilbur (Eds.). 1997. *Video-Mediated Communication*. Lawrence Erlbaum Associates, Mahwah, New Jersy.

[22] Ulta Frith. 1989. *Autism: explaining the enigma*. Blackwell Publishers, Oxford.

[23] Anthony Giddens. 1984. *The Constitution of Society*. University of California Press, Berkeley, CA.

[24] Kristen Gillespie-Lynch, Steven K. Kapp, Christina Shane-Simpson, David Shane Smith, and Ted Hutman. 2014. Intersections Between the Autism Spectrum and the Internet: Perceived Benefits and Preferred Functions of Computer-Mediated Communication. *Intellectual and Developmental Disabilities* 52, 6 (Dec. 2014), 456–469. https://doi.org/10.1352/1934-9556-52.6.456

[25] David M. Grayson and Andrew F. Monk. 2003. Are you looking at me? Eye contact and desktop video conferencing. *ACM Transactions on Computer-Human Interaction* 10, 3 (Sept. 2003), 221–243. https://doi.org/10.1145/937549.937552

[26] James J Gross and Robert W Levenson. 1997. Hiding feelings: the acute effects of inhibiting negative and positive emotion. *Journal of abnormal psychology* 106, 1 (1997), 95.

[27] Bernie Hogan. 2010. The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society* 30, 6 (Dec. 2010), 377–386. https://doi.org/10.1177/0270467610385893

[28] Laura Hull, K. V. Petrides, Carrie Allison, Paula Smith, Simon Baron-Cohen, Meng-Chuan Lai, and William Mandy. 2017. "Putting on My Best Normal": Social Camouflaging in Adults with Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders* 47, 8 (01 Aug 2017), 2519–2534. https://doi.org/10.1007/s10803-017-3166-5

[29] Lorcan Kenny, Caroline Hattersley, Bonnie Molins, Carole Buckley, Carol Povey, and Elizabeth Pellicano. 2015. Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism* 20, 4 (2015), 442–462. http://journals.sagepub.com/doi/abs/10.1177/1362361315588200

[30] Felipe Korzenny. 1978. A Theory of Electronic Propinquity: Mediated Communication in Organizations. *Communication Research* 5, 1 (Jan. 1978), 3–24. https://doi.org/10.1177/009365027800500101

[31] Rachel Loomes, Laura Hull, and William Polmear Locke Mandy. 2017. What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry* 56, 6 (2017), 466–474.

[32] Miriam Madsen, Rana el Kaliouby, Matthew Goodwin, and Rosalind Picard. 2008. Technology for just-in-time in-situ learning of facial affect for persons diagnosed with an autism spectrum disorder. In *Proceedings of the 10th International ACM SIGACCESS conference on Computers and Accessibility - ASSETS 2008*. ACM Press, Halifax, Nova Scotia, Canada, 19. https://doi.org/10.1145/1414471.1414477

[33] R. R. McCrae and O. P. John. 1992. An Introduction to the Five-Factor Model and its Application. *Journal of Personality* 60, 2 (June 1992), 175–215.

[34] Matthew Miles, A. Michael Huberman, and Michael Saldaña. 2013. *Qualitative Data Analysis: A Methods Sourcebook*. Sage Publications, Thousand Oaks, CA.

[35] Christine Miserandino. 2003. The Spoon Theory. https://butyoudontlooksick.com/articles/written-by-christine/the-spoon-theory/

[36] Meredith Ringel Morris, Andrew Begel, and Ben Wiedermann. 2015. Understanding the Challenges Faced by Neurodiverse Software Engineering Employees: Towards a More Inclusive and Productive Technical Workforce. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS '15)*. ACM, New York, NY, USA, 173–184. https://doi.org/10.1145/2700648.2809841

[37] Kenton O'Hara, Alison Black, and Matthew Lipson. 2006. Everyday practices with mobile video telephony. In *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*. ACM Press, Montréal, Québec, Canada, 871. https://doi.org/10.1145/1124772.1124900

[38] Mike Oliver. 2013. The social model of disability: thirty years on. *Disability & Society* 28, 7 (Oct. 2013), 1024–1026. https://doi.org/10.1080/09687599.2013.818773

[39] Michelle O'Reilly, Jessica Nina Lester, and Tom Muskett (Eds.). 2017. *A Practical Guide to Social Interaction Research in Autism Spectrum Disorders*. Palgrave Macmillan, London, UK.

[40] Rosalind W Picard. 2000. *Affective computing*. MIT press, Cambridge, MA.

[41] Amit Pinchevski and John Durham Peters. 2016. Autism and new media: Disability between technology and society. *New Media & Society* 18, 11 (Dec. 2016), 2507–2523. https://doi.org/10.1177/1461444815594441

[42] Gnanathusharan Rajendran and Peter Mitchell. 2007. Cognitive theories of autism. *Developmental Review* 27, 2 (2007), 224 – 260. https://doi.org/10.1016/j.dr.2007.02.001

[43] Kathryn E. Ringland, Christine T. Wolf, LouAnne E. Boyd, Mark S. Baldwin, and Gillian R. Hayes. 2016. Would You Be Mine: Appropriating Minecraft as an Assistive Technology for Youth with Autism. In *Proceedings of the 18th International ACM SIGACCESS conference on Computers and Accessibility - ASSETS 2016*. ACM Press, Reno, NV, 33–41. https://doi.org/10.1145/2982142.2982172

[44] Marwin Schmitt, Simon Gunkel, Pablo Cesar, and Dick Bulterman. 2014. Mitigating problems in video-mediated group discussions: Towards conversation aware video-conferencing systems. In *Proceedings of the 2014 workshop on Understanding and Modeling Multiparty, Multimodal Interactions - UM3I '14*. ACM Press, Istanbul, Turkey, 39–44. https://doi.org/10.1145/2666242.2666247

[45] Jeremiah Scholl, John McCarthy, and Rikard Harr. 2006. A comparison of chat and audio in media rich environments. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work - CSCW '06*. ACM Press,

Banff, Alberta, Canada, 323. https://doi.org/10.1145/1180875.1180925

[46] Leslie D. Setlock and Susan R. Fussell. 2010. What's it worth to you?: the costs and affordances of CMC tools to asian and american users. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10*. ACM Press, Savannah, Georgia, USA, 341. https://doi.org/10.1145/1718918.1718979

[47] Judy Singer. 1999. 'Why can't you be normal for once in your life?' From a 'problem with no name' to the emergence of a new category of difference. In *Disability Discourse*, Judy Singer and Sally French (Eds.). Open University Press, Philadelphia, PA, 59–67.

[48] Daniel P. Skorich, Tahlia B. Gash, Katie L. Stalker, Lidan Zheng, and S. Alexander Haslam. 2017. Exploring the Cognitive Foundations of the Shared Attention Mechanism: Evidence for a Relationship Between Self-Categorization and Shared Attention Across the Autism Spectrum. *Journal of Autism and Developmental Disorders* 47, 5 (01 May 2017), 1341–1353. https://doi.org/10.1007/s10803-017-3049-9

[49] D. Tannen. 2005. *Conversational Style: Analyzing Talk Among Friends*. Oxford University Press, New York.

[50] Silvan S Tomkins. 1984. Affect theory. In *Approaches to emotion*, Klaus Sherer and Paul Ekman (Eds.). L. Erlbaum Associates, Hillsdale, NJ, 163–195.

[51] Gina Venolia, John Tang, Ruy Cervantes, Sara Bly, George Robertson, Bongshin Lee, and Kori Inkpen. 2010. Embodied social proxy: mediating interpersonal connection in hub-and-satellite teams. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, Atlanta, Georgia, USA, 1049. https://doi.org/10.1145/1753326.1753482

[52] Joseph B. Walther. 1992. Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective. *Communication Research* 19, 1 (1992), 52–90. https://doi.org/10.1177/009365092019001003 arXiv:https://doi.org/10.1177/009365092019001003

[53] Joseph B. Walther. 1996. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research* 23, 1 (1996), 3–43.

[54] Joseph B. Walther, Brandon Van Der Heide, Artemio Ramirez, Judee K. Burgoon, and Jorge Peña. 2015. Interpersonal and Hyperpersonal Dimensions of Computer-Mediated Communication. In *The Handbook of the Psychology of Communication Technology*, S. Shyam Sundar (Ed.). John Wiley & Sons, Ltd, Chichester, UK, 1–22. https://doi.org/10.1002/9781118426456.ch1

[55] Peter Washington, Dennis Wall, Catalin Voss, Aaron Kline, Nick Haber, Jena Daniels, Azar Fazel, Titas De, Carl Feinstein, and Terry Winograd. 2017. SuperpowerGlass: A Wearable Aid for the At-Home Therapy of Children with Autism. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (Sept. 2017), 1–22. https://doi.org/10.1145/3130977

[56] Jurgen Wegge. 2006. Communication via Videoconference: Emotional and Cognitive Consequences of Affective Personality Dispositions, Seeing One's Own Picture, and Disturbing Events. *Human-Computer Interaction* 21, 3 (Sept. 2006), 273–318. https://doi.org/10.1207/s15327051hci2103_1

[57] Heinz Wimmer and Josef Perner. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 1 (1983), 103–128.

[58] Annuska Zolyomi, Anne Spencer Ross, Arpita Bhattacharya, Lauren Milne, and Sean A. Munson. 2018. Values, Identity, and Social Translucence: Neurodiverse Student Teams in Higher Education. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, Montreal QC, Canada, 1–13. https://doi.org/10.1145/3173574.3174073